养元素对另一种的导数)等于它们的价格之反比时,可取得资源或投入的最小成本组合.

$$\begin{cases} \dfrac{\partial P}{\partial N} = \dfrac{\partial Q}{\partial N}\Big/\left(\dfrac{\partial Q}{\partial P}\right) = -\dfrac{P_N}{P_P} \\ \dfrac{\partial K}{\partial N} = \dfrac{\partial Q}{\partial N}\Big/\left(\dfrac{\partial Q}{\partial K}\right) = -\dfrac{P_N}{P_K} \\ \dfrac{\partial K}{\partial P} = \dfrac{\partial Q}{\partial P}\Big/\left(\dfrac{\partial Q}{\partial K}\right) = -\dfrac{P_P}{P_K} \end{cases} \tag{7}$$

从(7)中即可求得一定产量下使成本最小的营养元素量.

五、模型优缺点与改进

模型最大优点在于对原始数据拟合时,采用多种方法进行,使之愈来愈完善,具有很高的拟合精度和适度性.在此基础上,对模型作进一步讨论便可得到一系列可靠而实用的信息.并且,所得结论与客观事实很好地吻合,从而进一步说明模型是合理的.

在实际问题中,产量受作物种类、植株密度、施肥量、气候条件等各种因素的作用.我们仅考虑了施肥量影响,但稍加修改便能适合不同情况,如:

1.考虑植株密度:在原有数据基础上,加上一组植株密度变化数据,用同样方法建立四元二次模型,并加以讨论.

2.土壤肥力影响:在实际环境中,每块地肥力不等,有高产田与低产田之分.将土壤肥力也当作影响作物产量的一个因子,同样可进行分析.

在模型建立中,还可进行异常值检验,将其删除或加权,重新拟合后讨论.

参 考 文 献

[1] 张乃生等,晋东南旱地玉米"产量——施肥"多元回归模型及其应用分析,数理统计与管理,1(1989),10—13.
[2] 约翰·内特(美)等,应用线性回归模型,中国统计出版社,1990.
[3] 北京大学概率统计系,SAS/STAT 软件"回归分析过程",1991.
[4] J·法朗士(英)等,农业中的数学模型,农业出版社,1991.
[5] 厄尔O·黑迪(美)等,农业生产函数,农业出版社,1991.

关于施肥效果分析问题的评注

项 可 风
(中国科学院系统科学研究所,北京 100080)

1992年全国大学生数学模型竞赛,北京赛区共有46个队参赛,其中有26个队选做《施肥效果分析》题.我参加了本题的阅卷工作,总的情况很不错,都抓住问题的实质.应用回归方法去建立模型,而后用统计方法分析施肥效果.北京师范大学数学系队,获得北京赛区的特等奖,本期发表该队喻梅,金青松,唐福明等三位同学的文章,作为本题最优秀的一份答卷,供读者参阅.

下面就本次竞赛中被普遍忽视的几个问题提出一点看法.本文所使用符号与数据可

参看喻梅等的文章。

一、多因素轮换法试验，不可能估计交互作用

在氮(N)、磷(P)、钾(K)三种施肥试验中，本题的设计是，当一个营养素的施肥量变化时，总将另二个营养素的施肥量保持在第7水平上，如对土豆产量关于N的施肥量做试验时，P与K的施肥量分别取为196kg/ha与372kg/ha。这样的设计称为因素轮换法。

现假定施肥量与产量间关系，可用三元二次多项表示：

$$E(Q)=b_0+b_NN+b_PP+b_KK+b_{NN}N^2+b_{PP}P^2+b_{KK}K^2 + b_{NP}NP+b_{NK}NK+b_{PK}PK, \tag{1}$$

式中最后三项为二个变元的交叉项，反应了因素间的交互效应。

记(n_0, p_0, k_0)为N、P、K施肥量的7水平值。若将坐标原点移到(n_0,p_0,k_0)点，模型(1)用新的坐标表示。模型方程是不变的。不难证明，在新的坐标表示下，(1)式所对应的设计矩阵X。它的最后三列全为零。相应地，回归系数估计的正规方程中，b_{NP}，b_{NK}，b_{PK}的系数也为零。也就是说，b_{NP}，b_{NK}，b_{PK}可任意给定。用统计术语来说，回归系数b_{NP}，b_{NK}，b_{PK}是不可估的。

但为什么用逐步回归方法对土豆产量建模时，却出现交叉项呢？

在对变量进行中心化和标准化后，虽然设计矩阵X的最后三列不为零列。但是可以证明它们可以分别用常数项，二个一次项所对应的三列线性表出。因此不能说它反应的是交互效应，更确切地说，交互效应与一次项效应混杂，不可能分离出来。这主要是因素轮换法设计的弊病所造成。

二、区组效应不可忽视

本试验在(n_0, p_0, k_0)点，重复了三次试验。以土豆为例，在肥量$n_0=259$，$p_0=196$，$k_0=372$(kg/ha)，点上三次试验的产量为43.15，41.26，38.43(t/ha)。平均值$\bar{y}=40.95$(t/ha)，标准差$\hat{\sigma}=2.38$，若用2σ原理，试验的30个值均落在$\bar{y}\pm2\sigma^2$区间(36.19，45.71)之中，这显然不合理。造成不合理的原因，在于我们假定三次重复产量的波动完全由随机误差产生。但实际上三次重复带有系统误差，这种系统误差主要来源于土壤肥力，生长期的管理措施等多种试验中的外界条件变化。在试验设计中，把在试验实施过程中，外界环境条件的差异所造成的系统偏差称为区组效应。在本题中，对应每种营养素的10次施肥试验，可以当作一个区组。区组内的10次试验可认为试验条件较为一致，而不同区组间，差别较大。设β_N，β_P，β_K分别为三个区组效应量。将其放入模型(1)式中，不失一般性，可假定

$$\beta_N+\beta_P+\beta_K=0, \tag{2}$$

这样我们就能得到唯一解。

在本问题中，由于交叉项不可估，模型(1)对三个变量来说是可分离的，可对三组数据，分别作成三个一元二次曲线。三个模型可得三个常数项，在常数项中，混杂着区组效应。每个常数项减去它们的平均值，即为(2)式中区组效应估计。在应用模型作施肥效果

分析时，区组效应要从模型中扣除，以提高使用的精确度。

三、试验设计的改进问题

在农作物产量的施肥试验中，因素之间常常存在着交互作用，这在设计试验时，应该把这点考虑进去．由于因素轮换法，不可能考察到交互效应，一般要用多因素的析因设计．假如，仅仅要拟合一个完全二次多项式模型，下面推荐一种响应曲面设计法．也叫二次复合设计．在三因素试验中，总共只要做15次试验，就可以用很简单方法，估计出10个回归系．现以土豆试验为例，其15个试验设计点为

试号＼因素	N	P	K	试号＼因素	N	P	K
1	179	146	272	9	356	196	372
2	179	146	472	10	161	196	372
3	179	246	272	11	259	257	372
4	179	246	472	12	259	135	372
5	339	146	272	13	259	196	494
6	339	146	472	14	259	196	250
7	339	246	272	15	259	196	372
8	339	246	472				

复合设计由三部分点组成，首先选择一个中心点，例如15号试验以原试验的7水平为中心点．而后以该点为中心，对每因素选一适当步长，再选二个水平值．例如上面设计，N、P、K的步长分别为

$$\Delta N = 80,\ \Delta P = 50,\ \Delta K = 100(\mathrm{kg/ha})$$

中心点下，上二个水平值为

$$N_0 \pm \Delta N = (179,339),\ P_0 \pm \Delta P = (146,246),\ K_0 \pm \Delta K = (272,472).$$

再按 $2^3 = 8$ 析因试验法排列8次试验，即上表中1—8号试验． 第三批试验是在以中心对称的坐标轴上的两个点，即上表的9—14号．水平计算按公式：

$$\mathrm{N} = N_0 \pm \sqrt{\frac{\sqrt{8 \times 15} - 8}{2}}\Delta N = (161,356),$$

$$\mathrm{P} = P_0 \pm \sqrt{\frac{\sqrt{8 \times 15} - 8}{2}}\Delta P = (135,257),$$

$$\mathrm{K} = K_0 \pm \sqrt{\frac{\sqrt{8 \times 15} - 8}{2}}\Delta K = (250,494).$$

这样每个因素共取了五个水平，但只从 $5^3 = 125$ 个网格点中选出15个试验点．由于设计的有规则性，保证回归系数计算也很简单．有兴趣读者可参看[1]中的第14章内容．

参 考 文 献

[1] 项可风、吴启光，试验设计与数据分析，上海科技出版社，1989年．